

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-078448

(43)Date of publication of application : 14.03.2000

(51)Int.Cl.

H04N 5/225

(21)Application number : 10-259190

(71)Applicant : CANON INC

(22)Date of filing : 31.08.1998

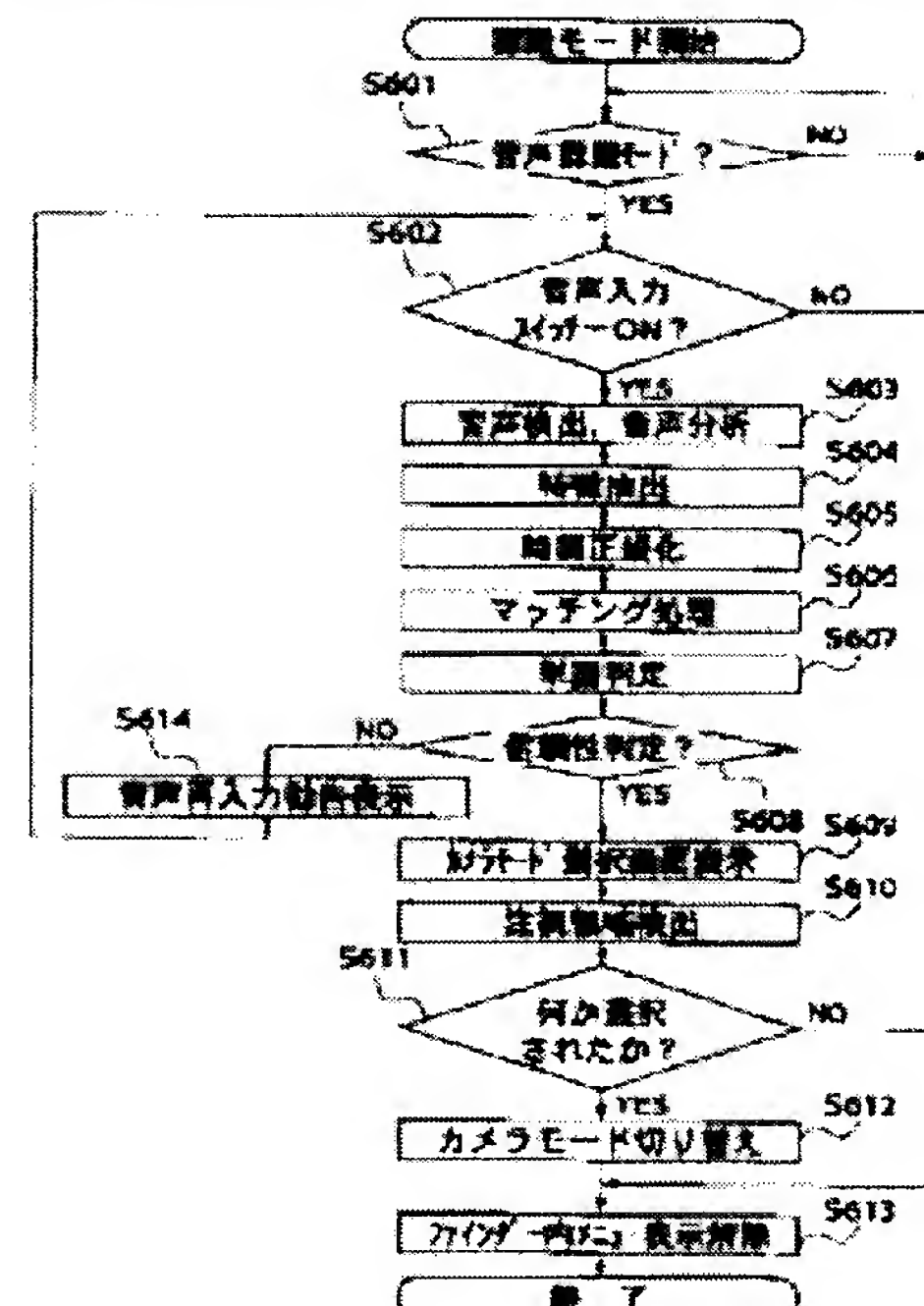
(72)Inventor : TAKEI HIROFUMI

(54) DEVICE AND METHOD FOR SELECTING CAMERA FUNCTION AND STORAGE MEDIUM

(57)Abstract:

PROBLEM TO BE SOLVED: To easily and surely set a function even in the case of a multifunctional voice control camera.

SOLUTION: The plural functions of a camera are hierarchically classified and the plural function names of a high-order hierarchy are displayed on a monitor picture display device. The voice of a photographer uttering one of plural function names of the high-order hierarchy displayed on the monitor picture display device is recognized (S603-S607), and the plural function names of the low-order hierarchy related to the function name shown by the recognized voice are displayed on the monitor picture display device (S609). In the picture showing the plural function names of the low-order hierarchy displayed on the monitor picture display device, a concerned point watched by the photographer is detected (S610) and based on the coordinates of the detected concerned point, the camera function is set (S612).



(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開2000-78448

(P2000-78448A)

(43)公開日 平成12年3月14日(2000.3.14)

(51)Int.Cl.⁷

識別記号

F I

テーマコード*(参考)

H 0 4 N 5/225

H 0 4 N 5/225

F 5 C 0 2 2

審査請求 未請求 請求項の数 8 F D (全 12 頁)

(21)出願番号 特願平10-259190

(22)出願日 平成10年8月31日(1998.8.31)

(71)出願人 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(72)発明者 竹井 浩文

東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

(74)代理人 100081880

弁理士 渡部 敏彦

Fターム(参考) 5C022 AA11 AB68 AC00 AC03 AC13

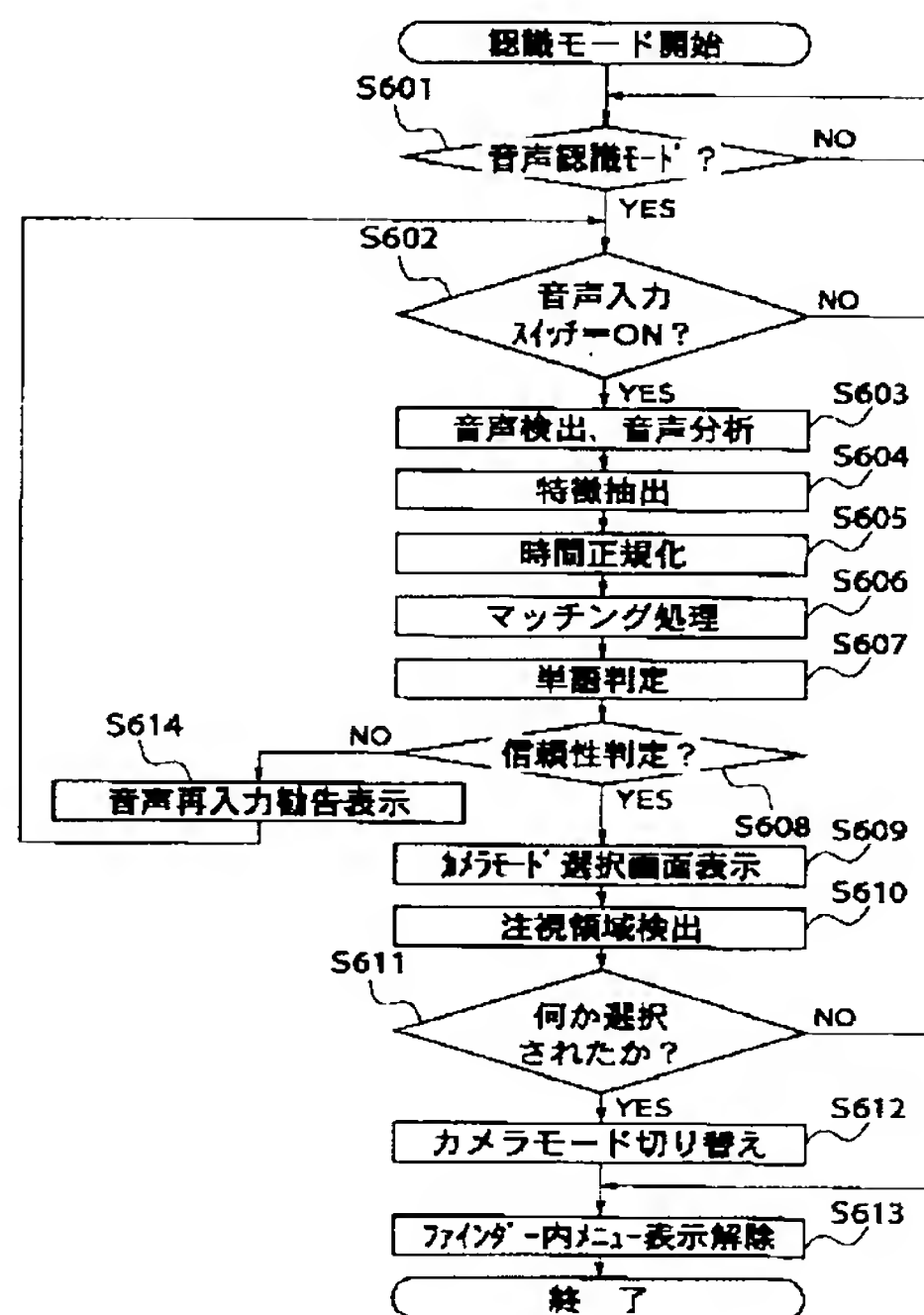
AC69 AC71 AC72 AC79

(54)【発明の名称】 カメラ機能選択装置、カメラ機能選択方法、及び記憶媒体

(57)【要約】

【課題】 多機能な音声制御カメラにおいても容易に確実に機能設定ができるようにする。

【解決手段】 カメラの持つ複数の機能を階層的に分類し、上位階層の複数機能名をモニタ画面表示装置に表示させる。モニタ画面表示装置に表示された上位階層の複数機能名の1つを発声した撮影者の音声を認識し(S603~S607)、認識された音声を示す機能名に関わる下位階層の複数機能名を、モニタ画面表示装置に表示させる(S609)。モニタ画面表示装置に表示された下位階層の複数機能名を示す画面において撮影者が注視した注視点を検出し(S610)、検出された注視点の座標に基づき、カメラ機能の設定を行う(S612)。



【特許請求の範囲】

【請求項 1】 モニタ画面表示装置を備えたカメラ機能選択装置において、
カメラの持つ複数の機能を階層的に分類し、上位階層の複数機能名をモニタ画面表示装置に表示させる表示手段と、
前記モニタ画面表示装置に表示された上位階層の複数機能名の 1 つを発声した撮影者の音声を認識する音声認識手段と、
前記音声認識手段によって認識された音声が示す機能名に関わる下位階層の複数機能名を、前記モニタ画面表示装置に表示させる音声制御手段と、
前記音声制御手段によって前記モニタ画面表示装置に表示された下位階層の複数機能名を示す画面において撮影者が注視した注視点を検出する視線検出手段と、
前記視線検出手段により検出された注視点の座標に基づき、カメラ機能の設定を行う機能設定手段とを有することを特徴とするカメラ機能選択装置。

【請求項 2】 所定時間に亘って、前記視線検出手段によって注視点が検出されない場合、前記モニタ画面表示装置に、下位階層の複数機能名を示す画面の表示を停止させる停止手段をさらに有することを特徴とする請求項 1 記載のカメラ機能選択装置。

【請求項 3】 モニタ画面表示装置を備えたカメラ機能選択装置において、
カメラの持つ複数の機能を階層的に分類し、上位階層の複数機能名をモニタ画面表示装置に表示させる表示手段と、
前記モニタ画面表示装置に表示された上位階層の複数機能名の 1 つを発声した撮影者の音声を認識する音声認識手段と、
前記音声認識手段によって認識された音声が示す機能名に関わる下位階層の複数機能名を、前記モニタ画面表示装置に表示させる音声制御手段と、
前記音声制御手段によって前記モニタ画面表示装置に表示された下位階層の複数機能名を示す画面において撮影者が触った位置を検出する接触位置検出手段と、
前記接触位置検出手段により検出された位置の座標に基づき、カメラ機能の設定を行う機能設定手段とを有することを特徴とするカメラ機能選択装置。

【請求項 4】 モニタ画面表示装置を備えたカメラ機能選択装置に適用されるカメラ機能選択方法において、
カメラの持つ複数の機能を階層的に分類し、上位階層の複数機能名をモニタ画面表示装置に表示させる表示ステップと、
前記モニタ画面表示装置に表示された上位階層の複数機能名の 1 つを発声した撮影者の音声を認識する音声認識ステップと、
前記音声認識ステップによって認識された音声が示す機能名に関わる下位階層の複数機能名を、前記モニタ画面

表示装置に表示させる音声制御ステップと、
前記音声制御ステップによって前記モニタ画面表示装置に表示された下位階層の複数機能名を示す画面において撮影者が注視した注視点を検出する視線検出ステップと、
前記視線検出ステップにより検出された注視点の座標に基づき、カメラ機能の設定を行う機能設定ステップとを有することを特徴とするカメラ機能選択方法。

【請求項 5】 所定時間に亘って、前記視線検出ステップによって注視点が検出されない場合、前記モニタ画面表示装置に、下位階層の複数機能名を示す画面の表示を停止させる停止ステップをさらに有することを特徴とする請求項 4 記載のカメラ機能選択方法。

【請求項 6】 モニタ画面表示装置を備えたカメラ機能選択装置に適用されるカメラ機能選択方法において、
カメラの持つ複数の機能を階層的に分類し、上位階層の複数機能名をモニタ画面表示装置に表示させる表示ステップと、
前記モニタ画面表示装置に表示された上位階層の複数機能名の 1 つを発声した撮影者の音声を認識する音声認識ステップと、
前記音声認識ステップによって認識された音声が示す機能名に関わる下位階層の複数機能名を、前記モニタ画面表示装置に表示させる音声制御ステップと、
前記音声制御ステップによって前記モニタ画面表示装置に表示された下位階層の複数機能名を示す画面において撮影者が触った位置を検出する接触位置検出ステップと、
前記接触位置検出ステップにより検出された位置の座標に基づき、カメラ機能の設定を行う機能設定ステップとを有することを特徴とするカメラ機能選択方法。

【請求項 7】 モニタ画面表示装置を備えたカメラ機能選択装置に適用されるカメラ機能選択方法をプログラムとして記憶した、コンピュータにより読み出し可能な記憶媒体において、
前記カメラ機能選択方法が、
カメラの持つ複数の機能を階層的に分類し、上位階層の複数機能名をモニタ画面表示装置に表示させる表示ステップと、
前記モニタ画面表示装置に表示された上位階層の複数機能名の 1 つを発声した撮影者の音声を認識する音声認識ステップと、
前記音声認識ステップによって認識された音声が示す機能名に関わる下位階層の複数機能名を、前記モニタ画面表示装置に表示させる音声制御ステップと、
前記音声制御ステップによって前記モニタ画面表示装置に表示された下位階層の複数機能名を示す画面において撮影者が注視した注視点を検出する視線検出ステップと、
前記視線検出ステップにより検出された注視点の座標に

基づき、カメラ機能の設定を行う機能設定ステップとを有することを特徴とする記憶媒体。

【請求項 8】 モニタ画面表示装置を備えたカメラ機能選択装置に適用されるカメラ機能選択方法をプログラムとして記憶した、コンピュータにより読み出し可能な記憶媒体において、

前記カメラ機能選択方法が、
カメラの持つ複数の機能を階層的に分類し、上位階層の複数機能名をモニタ画面表示装置に表示させる表示ステップと、

前記モニタ画面表示装置に表示された上位階層の複数機能名の 1 つを発声した撮影者の音声を認識する音声認識ステップと、

前記音声認識ステップによって認識された音声を示す機能名に関わる下位階層の複数機能名を、前記モニタ画面表示装置に表示させる音声制御ステップと、

前記音声制御ステップによって前記モニタ画面表示装置に表示された下位階層の複数機能名を示す画面において撮影者が触った位置を検出する接触位置検出ステップと、

前記接触位置検出ステップにより検出された位置の座標に基づき、カメラ機能の設定を行う機能設定ステップとを有することを特徴とする記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、カメラ機能選択装置、カメラ機能選択方法、及び記憶媒体に関し、詳しくは、撮影者の音声を認識し、その認識結果に応じてカメラの諸機能の中から所望の機能を選択するカメラ機能選択装置、当該カメラ機能選択装置に適用されるカメラ機能選択方法、及び当該カメラ機能選択方法を実行するプログラムを記憶した記憶媒体に関する。

【0002】

【従来の技術】最近のビデオカメラは高度に電子化され、小型なボディサイズにも拘わらず非常に多くの機能を備えることが可能となってきている。しかしながら、それに伴い、これらの機能を操作する為に電子ダイヤル、押し釦、スライドスイッチ等の操作部材が数多く用いられ、操作方法が判りづらくなっている。それとともに、限られたビデオカメラのサイズでは配置できる操作部材の数に限りがあるため、複数の操作部材を同時に押したり、順次階層的に操作するといった操作が導入されているが、結果的に、操作が複雑で面倒なものとなっている。

【0003】こうした不具合を解消する技術として、例えば特開昭64-56428号公報で提案される音声制御カメラがある。これによれば、操作者の音声を入力する音声入力手段と、入力された音声を認識する音声認識手段と、認識結果に対応する制御内容に基づいてカメラの機能を制御する制御手段とを有し、操作者の音声によ

って、絞り、シャッタ速度、動作モード等のカメラの機能を自由に設定できるようになっている。このような音声制御カメラにより複雑な操作をする必要がなくなり、機能の設定を簡便に行うことができるようになった。

【0004】

【発明が解決しようとする課題】しかしながら、こうした従来の音声制御カメラにおいて、多数存在するカメラの機能を操作者が音声で設定しようとした場合、設定する機能数が少ない場合は問題ないが、数が多くなるにつれて、どの機能をどの音声で設定したかを操作者が忘れてしまう場合がある。また多数の機能を階層構造にまとめておき、音声によって階層構造を順次辿り、目的の機能まで絞り込む、という方法も可能ではあるが、その場合、目的の機能に達するまでに何回も音声を発しなくてはならないので効率的でなく、また周囲の人に耳障りである、などの問題点があった。

【0005】本発明はこのような問題点に鑑みてなされたものであって、多機能な音声制御カメラにおいても容易に確実に機能設定ができる操作性の良いカメラ機能選択装置、カメラ機能選択方法、及び記憶媒体を提供することを目的とする。

【0006】

【課題を解決するための手段】上記目的を達成するために、請求項1記載の発明によれば、モニタ画面表示装置を備えたカメラ機能選択装置において、カメラの持つ複数の機能を階層的に分類し、上位階層の複数機能名をモニタ画面表示装置に表示させる表示手段と、前記モニタ画面表示装置に表示された上位階層の複数機能名の 1 つを発声した撮影者の音声を認識する音声認識手段と、前記音声認識手段によって認識された音声を示す機能名に関わる下位階層の複数機能名を、前記モニタ画面表示装置に表示させる音声制御手段と、前記音声制御手段によって前記モニタ画面表示装置に表示された下位階層の複数機能名を示す画面において撮影者が注視した注視点を検出する視線検出手段と、前記視線検出手段により検出された注視点の座標に基づき、カメラ機能の設定を行う機能設定手段とを有することを特徴とする。

【0007】また、請求項3記載の発明によれば、モニタ画面表示装置を備えたカメラ機能選択装置において、カメラの持つ複数の機能を階層的に分類し、上位階層の複数機能名をモニタ画面表示装置に表示させる表示手段と、前記モニタ画面表示装置に表示された上位階層の複数機能名の 1 つを発声した撮影者の音声を認識する音声認識手段と、前記音声認識手段によって認識された音声を示す機能名に関わる下位階層の複数機能名を、前記モニタ画面表示装置に表示させる音声制御手段と、前記音声制御手段によって前記モニタ画面表示装置に表示された下位階層の複数機能名を示す画面において撮影者が触った位置を検出する接触位置検出手段と、前記接触位置検出手段により検出された位置の座標に基づき、カメラ

機能の設定を行う機能設定手段とを有することを特徴とする。

【0008】さらに、請求項4記載の発明によれば、モニタ画面表示装置を備えたカメラ機能選択装置に適用されるカメラ機能選択方法において、カメラの持つ複数の機能を階層的に分類し、上位階層の複数機能名をモニタ画面表示装置に表示させる表示ステップと、前記モニタ画面表示装置に表示された上位階層の複数機能名の1つを発声した撮影者の音声を認識する音声認識ステップと、前記音声認識ステップによって認識された音声が表示機能名に関わる下位階層の複数機能名を、前記モニタ画面表示装置に表示させる音声制御ステップと、前記音声制御ステップによって前記モニタ画面表示装置に表示された下位階層の複数機能名を示す画面において撮影者が注視した注視点を検出する視線検出ステップと、前記視線検出ステップにより検出された注視点の座標に基づき、カメラ機能の設定を行う機能設定ステップとを有することを特徴とする。

【0009】また、請求項6記載の発明によれば、モニタ画面表示装置を備えたカメラ機能選択装置に適用されるカメラ機能選択方法において、カメラの持つ複数の機能を階層的に分類し、上位階層の複数機能名をモニタ画面表示装置に表示させる表示ステップと、前記モニタ画面表示装置に表示された上位階層の複数機能名の1つを発声した撮影者の音声を認識する音声認識ステップと、前記音声認識ステップによって認識された音声が表示機能名に関わる下位階層の複数機能名を、前記モニタ画面表示装置に表示させる音声制御ステップと、前記音声制御ステップによって前記モニタ画面表示装置に表示された下位階層の複数機能名を示す画面において撮影者が触った位置を検出する接触位置検出ステップと、前記接触位置検出ステップにより検出された位置の座標に基づき、カメラ機能の設定を行う機能設定ステップとを有することを特徴とする。

【0010】さらに、請求項7記載の発明によれば、モニタ画面表示装置を備えたカメラ機能選択装置に適用されるカメラ機能選択方法をプログラムとして記憶した、コンピュータにより読み出し可能な記憶媒体において、前記カメラ機能選択方法が、カメラの持つ複数の機能を階層的に分類し、上位階層の複数機能名をモニタ画面表示装置に表示させる表示ステップと、前記モニタ画面表示装置に表示された上位階層の複数機能名の1つを発声した撮影者の音声を認識する音声認識ステップと、前記音声認識ステップによって認識された音声が表示機能名に関わる下位階層の複数機能名を、前記モニタ画面表示装置に表示させる音声制御ステップと、前記音声制御ステップによって前記モニタ画面表示装置に表示された下位階層の複数機能名を示す画面において撮影者が注視した注視点を検出する視線検出ステップと、前記視線検出ステップにより検出された注視点の座標に基づき、カメ

ラ機能の設定を行う機能設定ステップとを有することを特徴とする。

【0011】また、請求項8記載の発明によれば、モニタ画面表示装置を備えたカメラ機能選択装置に適用されるカメラ機能選択方法をプログラムとして記憶した、コンピュータにより読み出し可能な記憶媒体において、前記カメラ機能選択方法が、カメラの持つ複数の機能を階層的に分類し、上位階層の複数機能名をモニタ画面表示装置に表示させる表示ステップと、前記モニタ画面表示装置に表示された上位階層の複数機能名の1つを発声した撮影者の音声を認識する音声認識ステップと、前記音声認識ステップによって認識された音声が表示機能名に関わる下位階層の複数機能名を、前記モニタ画面表示装置に表示させる音声制御ステップと、前記音声制御ステップによって前記モニタ画面表示装置に表示された下位階層の複数機能名を示す画面において撮影者が触った位置を検出する接触位置検出ステップと、前記接触位置検出ステップにより検出された位置の座標に基づき、カメラ機能の設定を行う機能設定ステップとを有することを特徴とする。

【0012】

【発明の実施の形態】以下、本発明の実施の形態を、図面を参照して説明する。

【0013】（第1の実施の形態）図1は本発明の第1の実施形態に係る、音声制御機能を備えたビデオカメラ（以下「カムコーダ」という）の構成を示すブロック図である。

【0014】図1において、ズームレンズ101からの被写体像は撮像素子102に結像し電気信号に変換され、映像信号処理回路103に入力される。撮像素子102は撮像素子駆動回路104により所望のシャッタスピードで駆動される。105はカムコーダ全体の制御、管理を行うメインCPUである。106はズームレンズ101にズーム動作を行わせるためのズーム駆動回路、107はフォーカス駆動回路、108は撮像素子102に人力する光量を調節する絞りを駆動する絞り駆動回路である。109はカムコーダの各種操作を行うキーユニットであり、手動でカムコーダの機能設定を行うメニューキー109aと、上下左右選択キー109bとを含む。キーユニット109において各キーを押し下げることによって発生した信号は、メインCPU105に読み込まれる。

【0015】110は映像信号を磁気テープなどに記録させるための記録装置である。111は、メインCPU105がファインダ画面114に表示したいカムコーダの各種状態情報をキャラクタ信号として変換作成するためのキャラクタ信号作成回路である。映像信号処理回路103からの映像信号とキャラクタ信号作成回路111からのキャラクタ信号は加算回路112で加算され、LCD表示回路113に入力され、ファインダ画面114

に映像として映し出される。撮影者は接眼レンズ115を通してファインダ画面114を観察できるようになっている。

【0016】116は、撮影者の注視点を検出する視線検出手段である。視線検出手段116は、撮影者の眼117に赤外光を照射する赤外発光ダイオード回路116aと、可視光を透過し赤外光を反射するダイクロイックミラー116bと、このダイクロイックミラー116bにて反射された赤外光を集光する集光レンズ116cと、この集光レンズ116cにより集光された赤外光を電気信号に変換する光電変換素子116dと、この光電変換素子116d上に結像した撮影者の眼117の像を基に、撮影者のファインダ画面114上の注視点を求める注視点検出回路116eとを具備している。ダイクロイックミラー116bは、可視光を透過するため、撮影者が接眼レンズ115を通してファインダ画面114を観察できるようになっている。また、ダイクロイックミラー116cは、赤外光を反射するため、赤外発光ダイオード回路116aからの赤外光が撮影者の眼117に照射され、反射して得られた反射像を集光レンズ116cへ向けて反射する。反射像は集光レンズ116cで集光されて光電変換素子116d上に像を結ぶ。注視点検出回路116eは、図示しないメモリを備え、光電変換素子116d上に結像した撮影者の眼117の像を基に、従来のアルゴリズムに従い撮影者のファインダ画面114上の注視点を求めるものである。このアルゴリズムとしては、例えば、特開平1-241511号公報、特開平2-32312号公報等に記載されたアルゴリズムを用いる。

【0017】118は音声認識手段である。音声認識手段118において、118aはマイクロコンピュータから構成され、主に音声認識処理をつかさどる音声認識回路である。撮影者の音声入力に従いマイク118bから出力された音声信号はプリアンプ118cに入力され、所定ゲインで増幅される。つぎにA/D変換器118dに送られてデジタル音声信号に変換され、音声認識回路118aに送られて音声認識処理が行われる。そして音声認識された結果や音声認識動作状況はデータバスによってメインCPU105に送信される。なお、音声認識回路118aからプリアンプ118cにフィードバック制御がかかるようになっており、音声認識回路118aに音声認識に適した音量が入力されるように、プリアンプ118cでは、いわゆるオートゲインコントロール（AGC）が行われる。

【0018】118eは、音声認識回路118aが音声認識処理を行う際のワーキングメモリーとして設けられたRAMであり、予め撮影者の音声の音響的特徴が格納される。SW-Mは音声モードスイッチであり、音声入力機能をオフするポジション、音声認識動作を行う音声認識モードのポジション、及び撮影者の音声を予め登録

する音声登録モードのポジションのいずれかを選択するために使用される。SW-Vは、撮影者が発声する音声を入力させる際のトリガを発生する音声入力スイッチである。

【0019】一般的に音声認識装置は、撮影者を限定する特定話者用と、操作者を限定せずに誰の声でも認識する不特定話者用とに分類される。特定話者用は、特定の撮影者に認識系を設定することができるため、装置の負荷が軽くなるとともに高い認識精度が期待でき、また認識精度が言語にも影響されにくい特性を持っている。しかし認識すべき音声を撮影者に予め発声させ、登録しておくという操作が必ず必要となる。一方、不特定話者用はそうした登録作業が不要であり、すぐに音声認識を実行できる簡便性はあるが、認識精度が低い。不特定話者用において認識精度を上げる為には、演算装置、メモリとも大規模なシステムが必要となってくる。

【0020】ところで、一般用ビデオカメラでは、撮影者はほとんどの場合一個人に限定され、装置は小型で低コストであることが求められる。そうした点を考慮すると、音声認識装置には、特定話者用でかつ特定音声を対象とする方式が適しているといえる。このような背景から本実施の形態で扱う、音声制御機能を備えたビデオカメラ（カムコーダ）も特定話者仕様になっている。

【0021】ここで、音声認識回路118aが行う音声認識処理について説明する。

【0022】一般的に音声認識の過程は、音声を、認識に役に立つなるべく少数の特徴パラメータで表すための特徴抽出部と、その特徴パラメータによって表された音声に相当する語彙を判定する判別部に分けることができる。これら認識技術については現在数々の研究がなされているが、代表的な手法として、認識の対象となる限られた数の単語について音声の標準パターンを作成し、それと入力音声との一致度を検出して判定するパターンマッチング方式がある。本実施の形態では、この方式を採用する。

【0023】図2は、パターンマッチング方式の認識処理の手順を示すフローチャートである。

【0024】特定話者から発声された音声データはステップS201にて、バンドパスフィルタ分析等の音声分析により、分析パラメータベクトルの時系列に変換されるとともに、音声の振幅パターン等から単語の開始点及び終了点が決定され、単語の切り出しが行われる。次にステップS202にて特徴点抽出が行われ、音声認識に有効な少数の特徴パラメータが抽出される。具体的には、ステップS201で得られたスペクトルのローカルピークを検出し、これらのみを2値化抽出する。これによりデータ圧縮が行われる。次にステップS203にて、線形又は非線形の時間正規化処理が行われ、音声パターンが生成される。前述のように特定話者対応の場合は予め撮影者の音声データを参照音声パターンとして登

録する必要がある、この登録を行なう登録モードに設定されているときには（S204でYES）、ステップS203で生成された音声パターンをメモリ118eに記憶させる（S206）。また、操作者（撮影者）の音声認識してカムコーダの機能を制御する認識モードに設定されているときには（S204でNO、S205でYES）、入力音声と参照音声パターンとのマッチング計算を行なうマッチング処理が行なわれる（S207）。マッチング計算は、時間正規化された参照音声パターンベクトルと入力音声パターンベクトルとの距離計算として行われる。最後にステップS208にて、登録された各参照音声パターンとの距離の中で最小のものが認識された単語として判定される。

【0025】次に第1の実施形態の具体的な動作について説明する。

【0026】図3は、認識すべき撮影者の音声を予め登録する「音声登録モード」における動作を示すフローチャートである。

【0027】音声認識回路118aは音声モードスイッチSW-Mの状態を検知し、スイッチが登録のポジションに位置していると「音声登録モード」であることをメインCPU105に通知する（S301）。するとメインCPU105はステップS302にて、機能設定メニューの表示から撮影者にカムコーダのモード設定を行わせる。

【0028】図4は、ファインダ画面114に表示されるモード設定画面を示す図であり、図4（A）、図4（B-1）、図4（B-2）は、手動による機能設定の一例としてAEモードが選択される様子を示すものである。まずカムコーダ本体のメニューキー109aを押下することでファインダ画面114の表示状態は図4

（A）に示すようになり、カメラ撮影時に設定できる機能の項目が多数表示される。ここでさらに上下左右選択キー109bの下矢印キーを押下することで図4（B-1）で示すようにカーソルを所望の機能項目に設定する（この場合はAEモードに設定）。次に上下左右選択キー109bの右矢印キーを押下することで、図4（B-2）のように、さらに細かなAEモードの項目が設定できる画面に切り替わる。こうして設定が終了した後、モード設定画面はそのまま表示が維持される（S303）。

【0029】次にステップS304にて、音声入力スイッチSW-Vが押されているか否かを音声認識回路118aが検出する。ステップS304にて、音声入力スイッチSW-Vが押されていると認識されれば、ステップS305に進み、音声認識回路118aはメインCPU105から設定内容を読み込むとともに、音声検出及び音声分析を行う。すなわち、表示されている選択モードと、音声認識させる撮影者の入力音声とを対応させて登録させるべく、撮影者にモード名を発声させる。この例

ではAEモード選択画面〔図4（B-2）〕に設定されているので「エーイー」という具合に撮影者が発声する。また図4（C-1）、図4（C-2）はシャッタースピードモードを選択する様子を示すが、こうしたシャッタースピードモード選択画面であれば、「シャッタ」と撮影者が発声する。なお、設定画面に書かれた語彙だけでなく任意の語彙を登録時に発声することで独自の音声を登録することもできる。

【0030】次のステップS305での音声検出及び音声分析、ステップS306での特徴抽出、ステップS307での時間正規化は、図2のステップS201～S203で説明した処理と同じである。

【0031】次にステップS308にて、音声パターンの信頼性判定が行なわれる。すなわち、生成された音声パターンが、参照パターンとして登録するのに値するレベルに達しているか否かを判定する。信頼性が不十分であると判定されるとステップS311に進み、登録が不可であり、再度登録動作を行う必要があるとして再入力を勧告する表示を行なう。これはファインダ画面114に図5（A）のように警告表示がされる。そしてこの勧告表示を所定時間行い、ステップS304に戻り、再度音声入力スイッチSW-Vが押されるのを待つ。信頼性が十分と判断されるとステップS309、S310に進み、音声パターン記憶用に設けられたメモリ118eに参照音声パターンを記憶させ、登録動作が完了する。

【0032】次に音声認識をカムコーダに行わせる「音声認識モード」について説明する。

【0033】図6は音声認識動作を示すフローチャートである。

【0034】音声認識回路118aは音声モードスイッチSW-Mの状態を検知し、スイッチが認識のポジションに位置していると「音声認識モード」であることをメインCPU105に通知する（S601）。次にステップS602にて、音声認識回路118aは音声入力スイッチSW-Vが押されているか否かを検出する。押されていないければステップS601に戻り、押されていればステップS603に進む。ステップS603では音声認識回路118aによる音声認識動作がスタートするとともに、音声認識回路118aはメインCPU105に通知し、他の操作スイッチの人力を受け付けないようにさせる。撮影者が、音声入力スイッチSW-Vを押した後、予め登録されている語彙の何れかを発声すると、音声検出、音声分析ステップS603、特徴抽出ステップS604、時間正規化ステップS605、マッチング処理ステップS606、単語判定ステップS607と進み、音声認識回路118aは一連の音声認識処理を行なう。

【0035】次にステップS608に進み、音声認識の信頼性判定が行なわれる。つまり音声認識回路118aは、入力された音声パターンと、認識された参照音声パ

10

20

30

40

50

ターンとの距離が所定の基準値よりも小さいか否かを判断する。大きければ認識信頼性がないと判断してステップS614に進む。ステップS614では、再度入力動作を行うようにファインダ画面114に、図5(B)に示すような勧告表示を行う。距離が小さければ認識信頼性が充分と判断し、ステップS609に進む。ステップS609では、音声認識回路118aがメインCPU105に認識結果を送信し、メインCPU105は認識結果に対応する撮影モードにカムコードの設定を切り換え、認識結果に対応するモード表示をファインダ画面114に表示する。

【0036】その後、ステップS610にて、注視点検出回路116eがファインダ画面114内の注視領域の検出を行う。この検出については、図7、8を参照して後述する。検出の結果、撮影者が注視により特定の領域を選択していれば(S611でYES)、注視している領域を示す情報がメインCPU105に転送され、ステップS612にて、選択された領域が相当する内容の撮影モードに設定される。その後、ファインダ画面114内のメニュー表示は解除され、ファインダ画面114は

通常画像表示を行う(S613)。

【0037】図7、8は、注視点検出回路116e内に含まれるマイクロコンピュータが実行する注視領域の検出処理の手順を示すフローチャートである。

【0038】検出処理に先立って予め、ファインダ画面114に表示される各機能表示画面における領域位置の座標が、注視点検出回路116e内のメモリにそれぞれ記憶される。

【0039】図9は、ファインダ画面114に表示される、例えばAEモードの機能表示画面における領域位置を示す図である。ファインダ画面114の一番上の項目から順に領域1～6として、注視点検出回路116e内のメモリに記憶される。

【0040】まず、ステップS701にて、タイマを所定の時間に設定し、注視領域の演算を行う最大時間を決める。次にステップS702にて注視座標の検出を行う。

【0041】ここで、変数a、b、c、d、e、fを、撮影者の注視点が領域1～6にそれぞれ一致した回数を示す変数とする。以下に、撮影者が図9の領域1を注視した場合を例にとって説明する。

【0042】ステップS703で、注視座標が領域1に略一致したと判断されると注視点検出回路116e内のマイクロコンピュータは、変数a以外の変数b、c、d、e、fを0にリセットし(S704)、変数aが所定の回数(例えば5回)以上か否かを判断する(S705)。少なければ領域1の部分に表示している文字を点滅させるようにメインCPU105に通知する(S706)。そして変数aに1を加える(S707)。ステップS705の処理で5以上と判断されればステップS7

08にて、領域1の部分に表示してある文字(この場合は「全自動」)を白抜き表示にし、ステップS709にて、領域1の部分が注視により選択されたことをメインCPU105に伝える。

【0043】領域1と注視座標とが一致しても、一致した回数が5回よりも少ないうちに注視座標が領域1から外れれば、変数aは0にリセットされる(S711、S718、S725、S732、S739、S745)。

【0044】その他の領域2～6を注視した場合についても同様の処理を行う。またS701で設定した所定時間内に、領域1～6のいずれの領域も注視により選択されなかった場合はステップS746にて処理を終える。

【0045】次に、こうしたカムコードを撮影者が使用して撮影を行う際の撮影者の動作例を説明する。

【0046】例えば、撮影者が音声入力スイッチSW-Vを押し、「エーイー」と発声すると、ファインダ画面114には、図10(A)に示す画面が表示される。その後撮影者は画面内に表示されたAEモード群の中から所望の項目領域を注視する。この例では「絞り優先」を注視したとする。すると図10(B)に示すように、注視した「絞り優先」の項目領域が反転すると共に、カムコードは絞り優先のAEモードに設定される。また例えば、シャッタースピードを変更したい場合は、撮影者が音声入力スイッチSW-Vを押し「シャッタ」と発声すると、ファインダ画面114には、図11(A)に示す画面が表示される。その後、撮影者は画面内に表示されたシャッタースピードの中で所望の値を注視する。この例では「1/1000」を注視したとする。すると図10(B)のように注視した値の項目「1/1000」が反転すると共に、カムコードのシャッタースピードが1/1000秒に設定される。ここで発声した後にファインダ画面114のどの項目も注視しなかった場合、図7のステップS701でタイマに設定された時間が経過した後に機能選択画面の表示が解除され、通常の映像表示画面に切り替わる。

【0047】このように、従来はモード設定を、多数の手動スイッチを使用して複雑な操作により時間をかけて行ったり、また譬え音声制御方式のカムコードであっても音声を何回も発声して設定しなければならなかったものが、第1の実施の形態では、まず音声により大まかな機能群を選択し、さらに視線によりこまかな機能を選択するという組み合わせで、容易にカムコードの各種機能設定ができるようになる。

【0048】また、第1の実施の形態によれば、撮影者の音声の登録を通常の手動設定状態と同じ手順で行っている。そのため、撮影者にとって登録操作が馴染みやすく、すぐに理解できるとともに、登録の際に機能と音声とを対応させるための特別な操作部材や表示手段を必要としない。

【0049】(第2の実施の形態)図12は、第2の実

施形態に係るカムコーダの構成を示すブロック図である。第2の実施形態の構成は第1の実施形態と基本的に同じであるので、同一構成部分には同一の参照符号を付して、その説明を省略する。

【0050】第2の実施形態では、第1の実施形態における視線検出手段116をなくし、大型LCD表示回路113bと、LCD画面114bと、LCD画面114bの表面上で指が触った位置を検出するタッチ位置センサ201と、タッチ位置センサ201からの情報によりタッチした領域を検出するタッチ位置検出回路202とを新たに設ける。そして、第1の実施形態における図6のステップS610で実行される注視領域検出処理に代わって、撮影者が指でタッチした領域の情報をメインCPU105に転送するようにする。

【0051】これにより、第1の実施形態と同様に、まず音声により大まかな機能群を選択し、つぎにLCD画面114bに表示された詳細な項目領域を指で触ることによって機能を選択する。かくして、第1の実施形態と同様に容易にカムコーダの各種機能設定ができるようになる。

【0052】なお、本発明を、複数の機器から構成されるシステムに適用しても、あるいは1つの機器からなる装置に適用してもよい。

【0053】また、前述した各実施形態の機能を実現するソフトウェアのプログラムコードを記憶した記憶媒体を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記憶媒体に格納されたプログラムコードを読み出して実行することによっても、本発明が達成されることは言うまでもない。

【0054】この場合、記憶媒体から読み出されたプログラムコード自体が、前述の各実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体が本発明を構成することになる。

【0055】プログラムコードを供給するための記憶媒体として、例えば、フロッピーディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、CD-R、磁気テープ、不揮発性のメモ리카ード、ROMなどを用いることができる。

【0056】また、コンピュータが読み出したプログラムコードを実行することにより、前述した各実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているOSなどが実際の処理の一部または全部を行い、その処理によって前述した各実施形態の機能が実現される場合も、本発明に含まれることは言うまでもない。

【0057】さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに

備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した各実施形態の機能が実現される場合も、本発明に含まれることは言うまでもない。

【0058】

【発明の効果】以上詳述したように請求項1、請求項4または請求項7記載の発明によれば、カメラの持つ複数の機能を階層的に分類し、上位階層の複数機能名をモニタ画面表示装置に表示させる。モニタ画面表示装置に表示された上位階層の複数機能名の1つを発声した撮影者の音声を認識し、認識された音声を示す機能名に関わる下位階層の複数機能名を、モニタ画面表示装置に表示させる。モニタ画面表示装置に表示された下位階層の複数機能名を示す画面において撮影者が注視した注視点を検出し、検出された注視点の座標に基づき、カメラ機能の設定を行う。

【0059】このように、音声認識と視線検出とを組み合わせることで使用することにより、多機能な音声制御カメラにおいても容易に確実に機能設定ができる。

【0060】また、請求項3、請求項6または請求項8記載の発明によれば、カメラの持つ複数の機能を階層的に分類し、上位階層の複数機能名をモニタ画面表示装置に表示させる。モニタ画面表示装置に表示された上位階層の複数機能名の1つを発声した撮影者の音声を認識し、認識された音声を示す機能名に関わる下位階層の複数機能名を、モニタ画面表示装置に表示させる。モニタ画面表示装置に表示された下位階層の複数機能名を示す画面において撮影者が接触した位置を検出し、検出された位置の座標に基づき、カメラ機能の設定を行う。

【0061】このように、音声認識と接触位置検出とを組み合わせることで使用することにより、多機能な音声制御カメラにおいても容易に確実に機能設定ができる。

【図面の簡単な説明】

【図1】第1の実施形態に係る、音声制御機能を備えたビデオカメラの構成を示すブロック図である。

【図2】パターンマッチング方式の認識処理の手順を示すフローチャートである。

【図3】認識すべき撮影者の音声を予め登録する「登録モード」における動作を示すフローチャートである。

【図4】ファインダ画面に表示されるモード設定画面を示す図である。

【図5】（A）は音声登録を失敗した場合のファインダ画面の警告表示例を示す図であり、（B）は音声認識を失敗した場合のファインダ画面の警告表示例を示す図である。

【図6】音声認識動作を示すフローチャートである。

【図7】注視点検出回路内に含まれるマイクロコンピュータが行う注視領域の検出処理の手順を示すフローチャートである。

【図8】注視点検出回路内に含まれるマイクロコンピュ

10

20

30

40

50

ータが行う注視領域の検出処理の手順を示すフローチャートである。

【図9】ファインダ画面に表示される、AEモードの機能表示画面における領域位置を示す図である。

【図10】AEモードを設定する場合のファインダ画面の表示例を示す図である。

【図11】シャッタースピードを設定する場合のファインダ画面の表示例を示す図である。

【図12】第2の実施形態に係るカムコードの構成を示すブロック図である。

【符号の説明】

105 メインCPU（表示手段、音声制御手段、機能*

*設定手段）

109 キーユニット

111 キャラクタ信号作成回路

113 LCD表示回路

114 ファインダ画面（モニタ画面表示装置）

115 接眼レンズ

116 視線検出手段

116e 注視点検出回路、

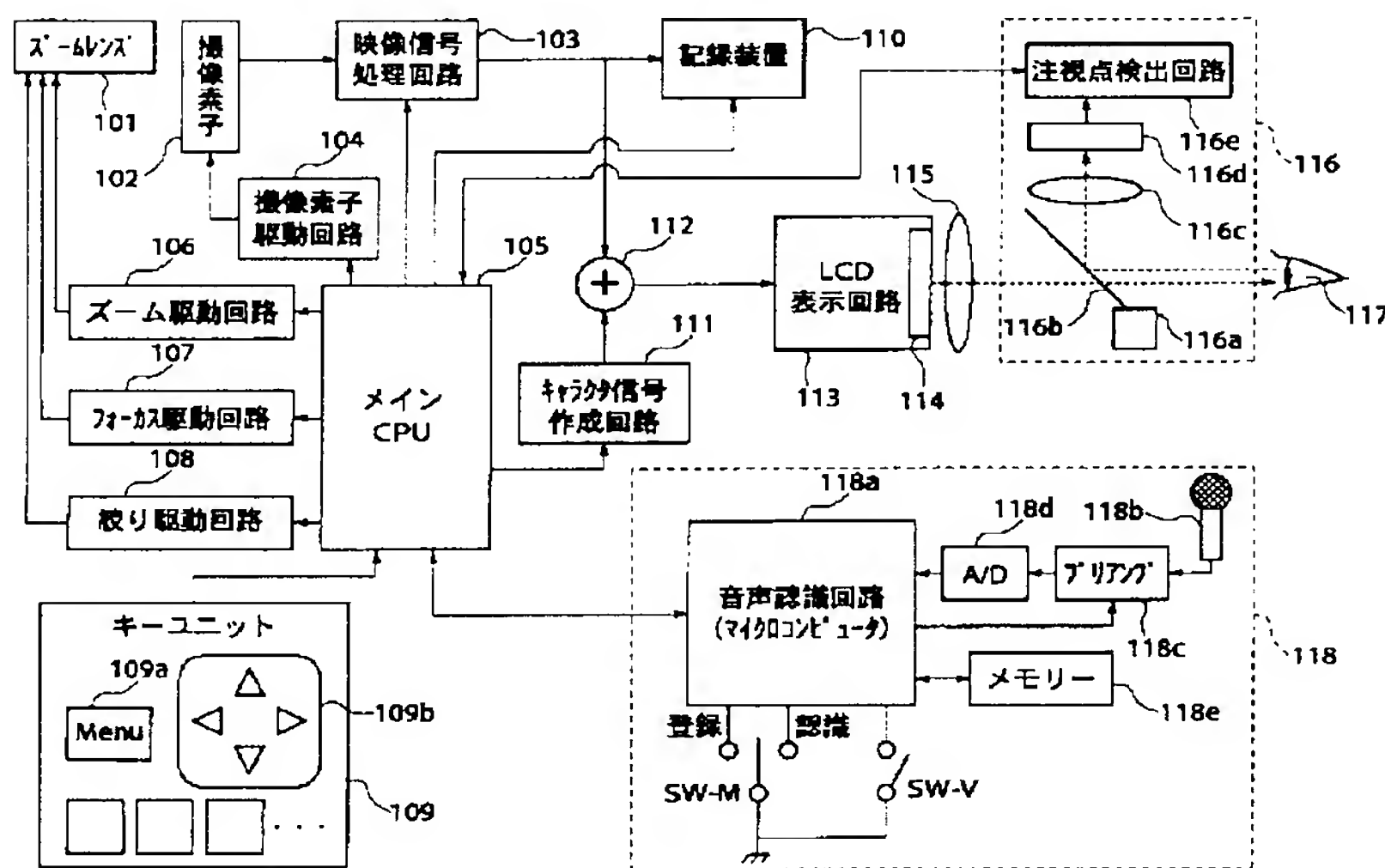
117 撮影者の眼

10 118 音声認識手段

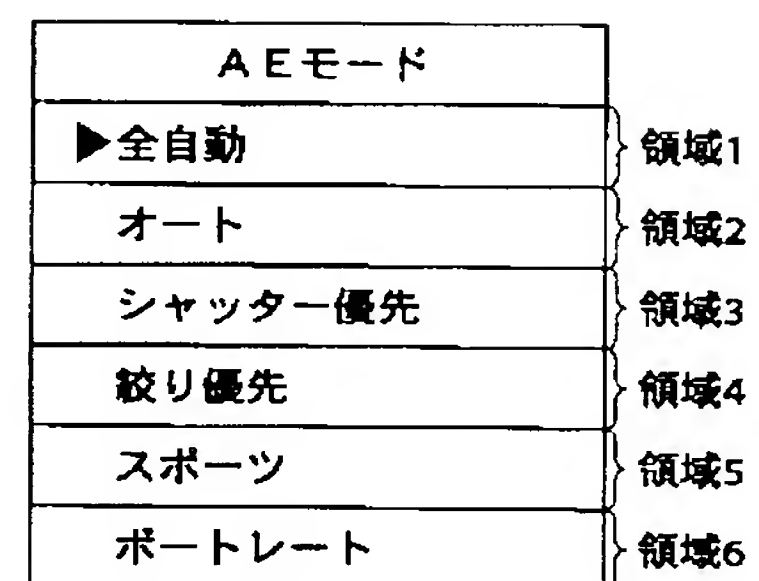
118a 音声認識回路

SW-V 音声入力スイッチ

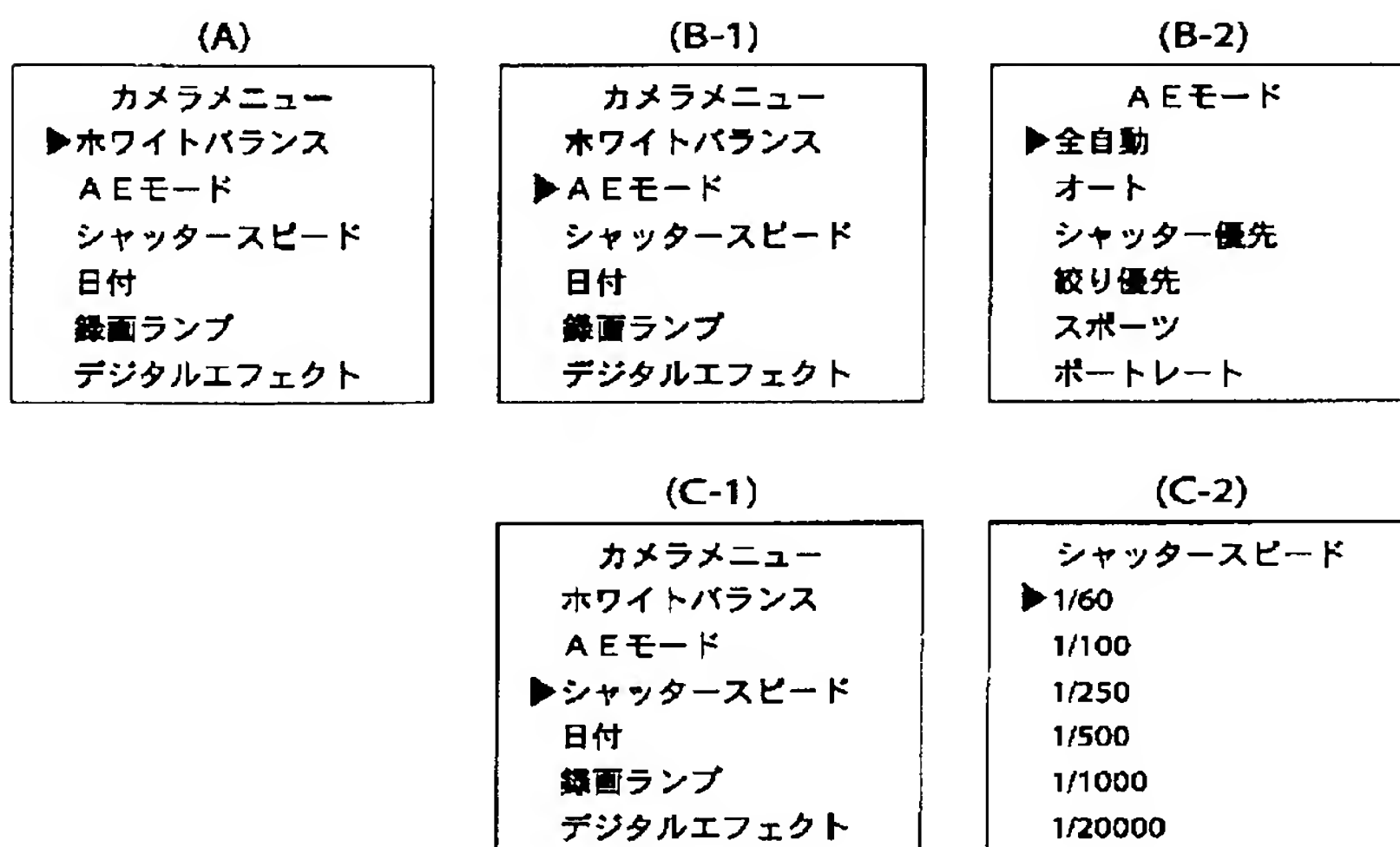
【図1】



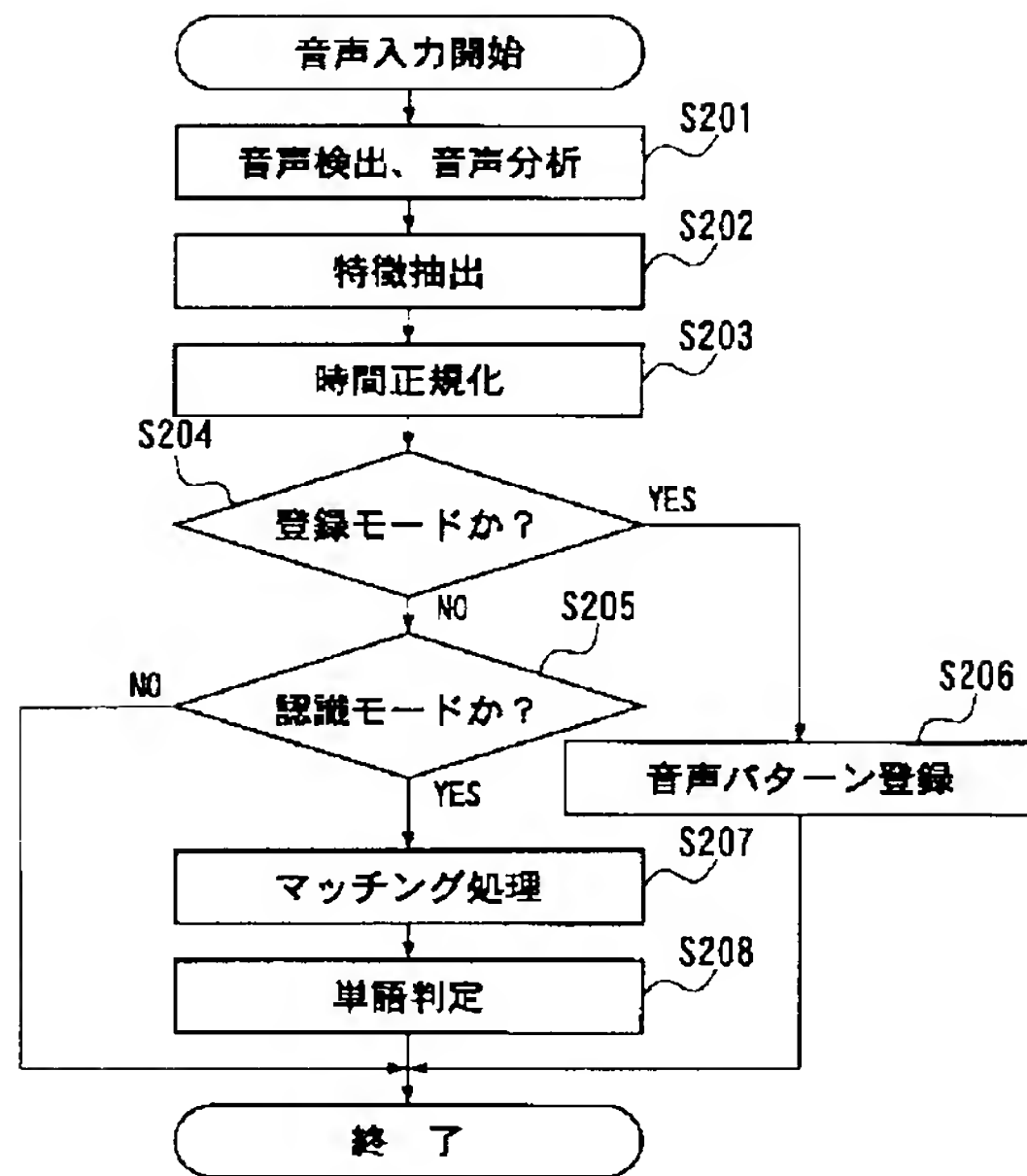
【図9】



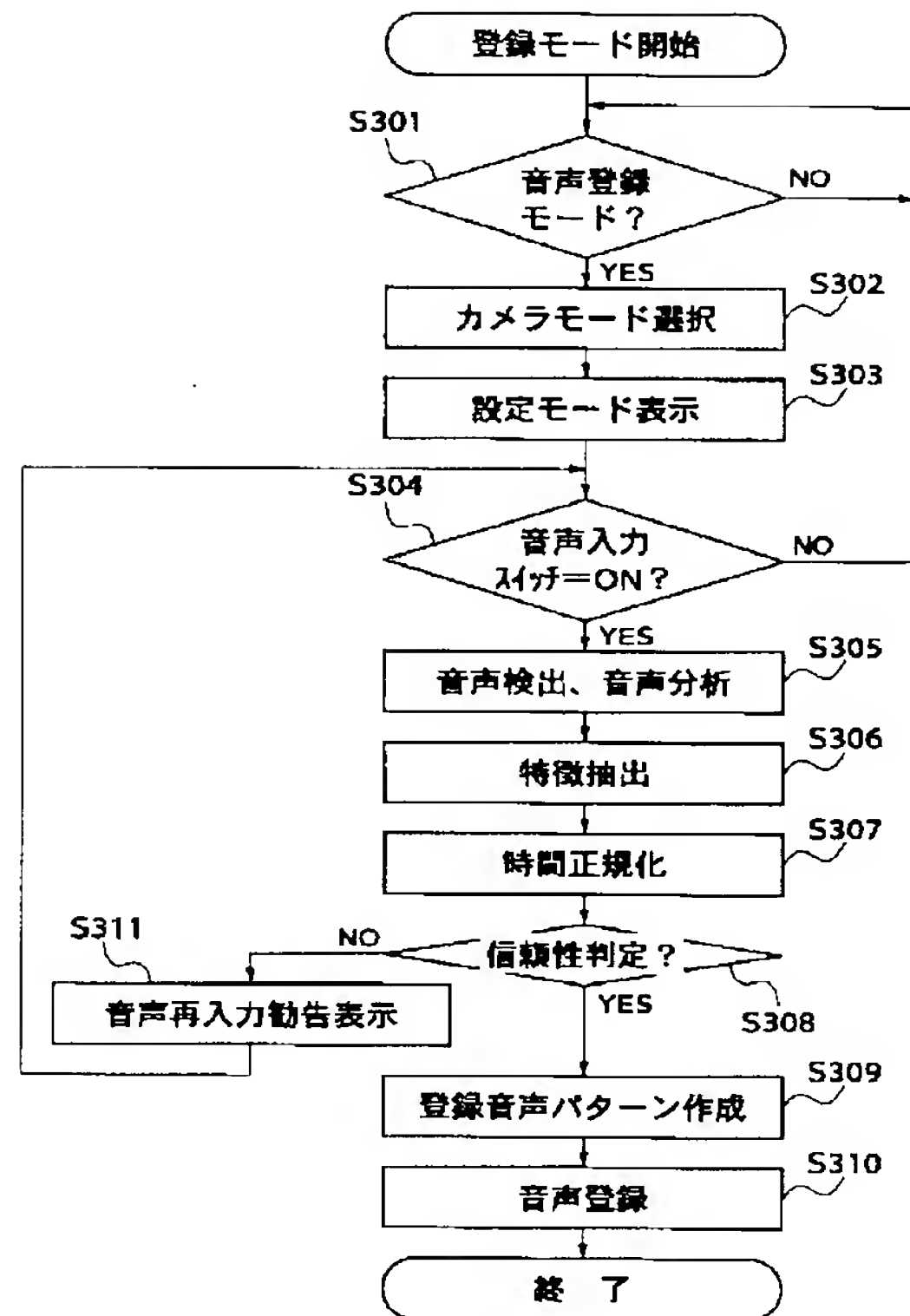
【図4】



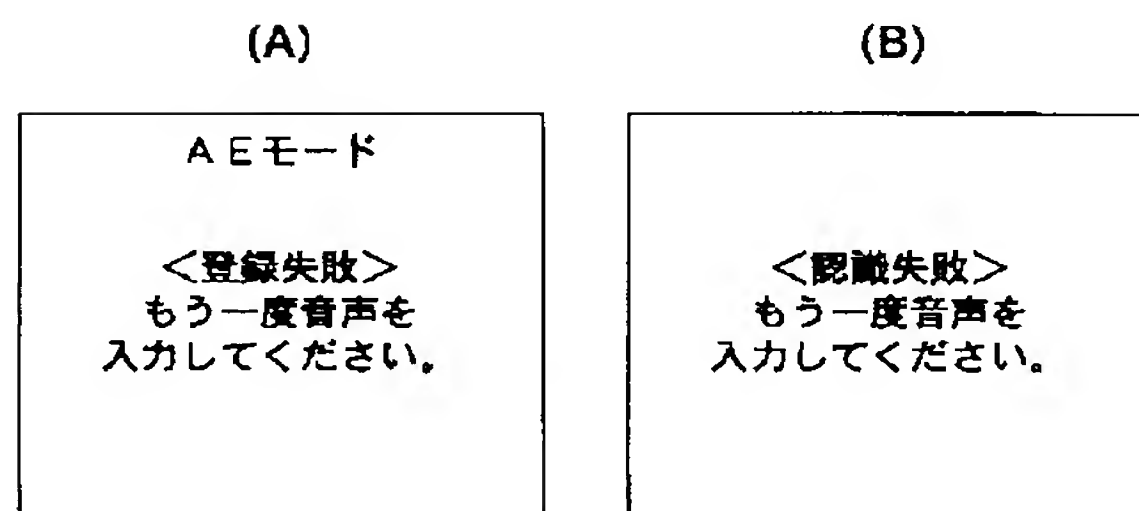
【図2】



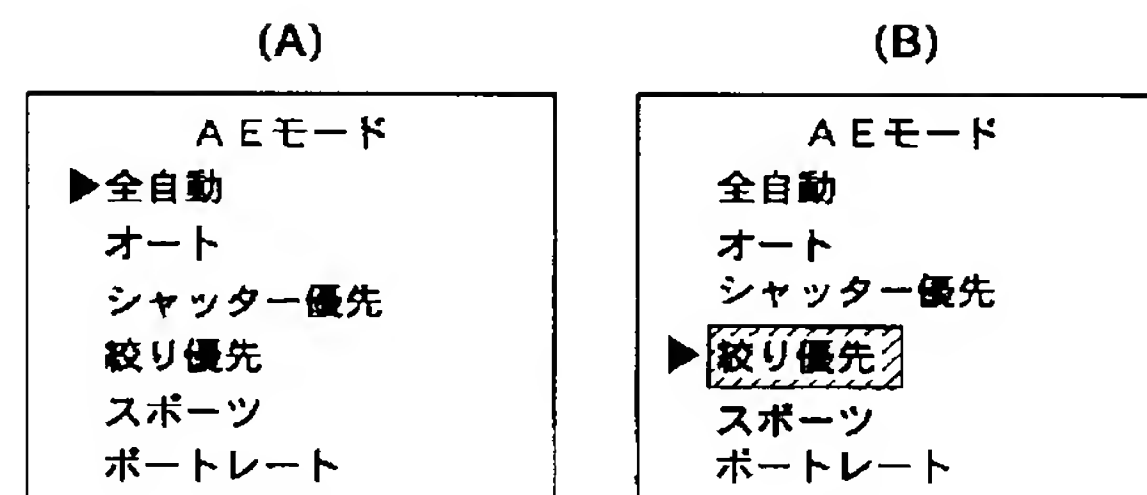
【図3】



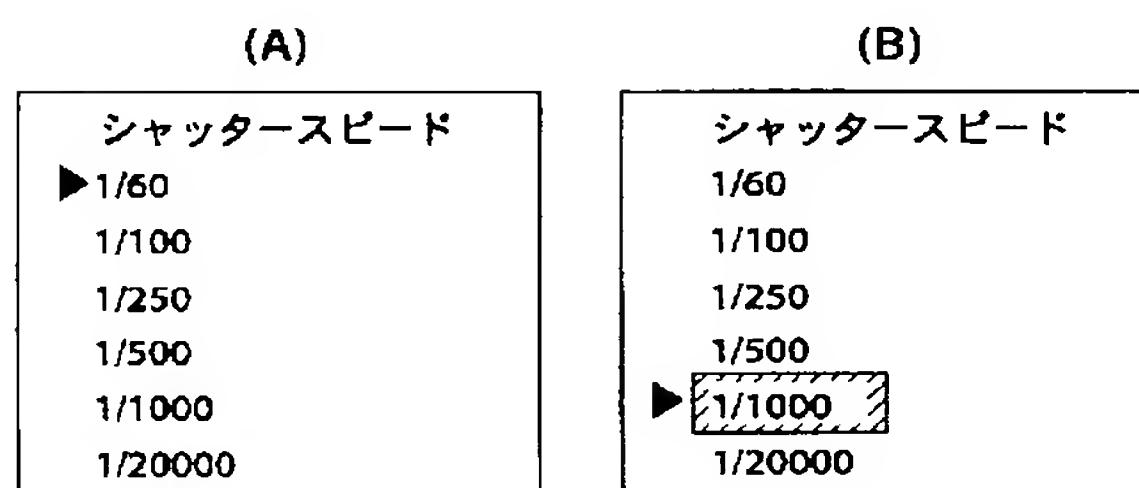
【図5】



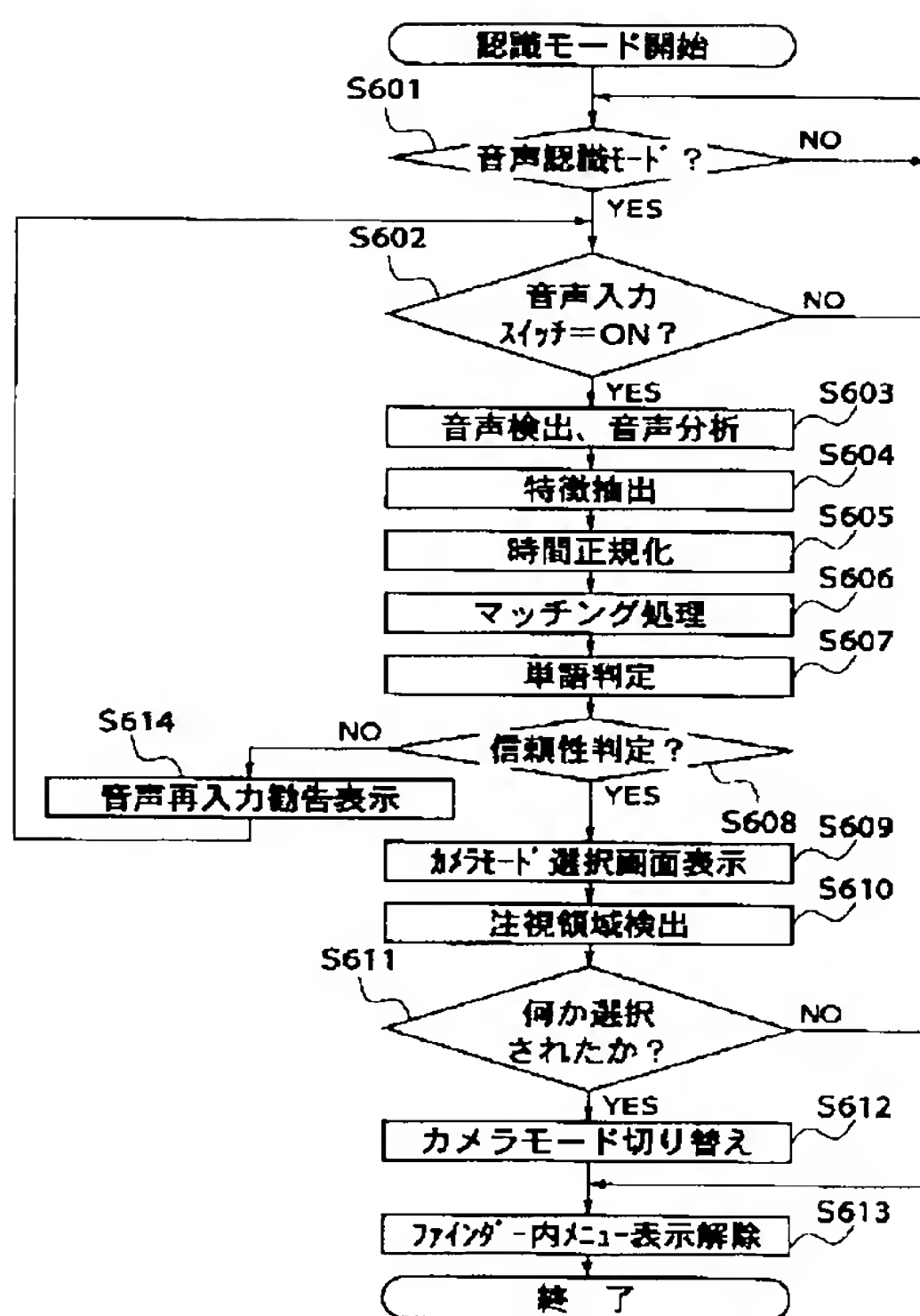
【図10】



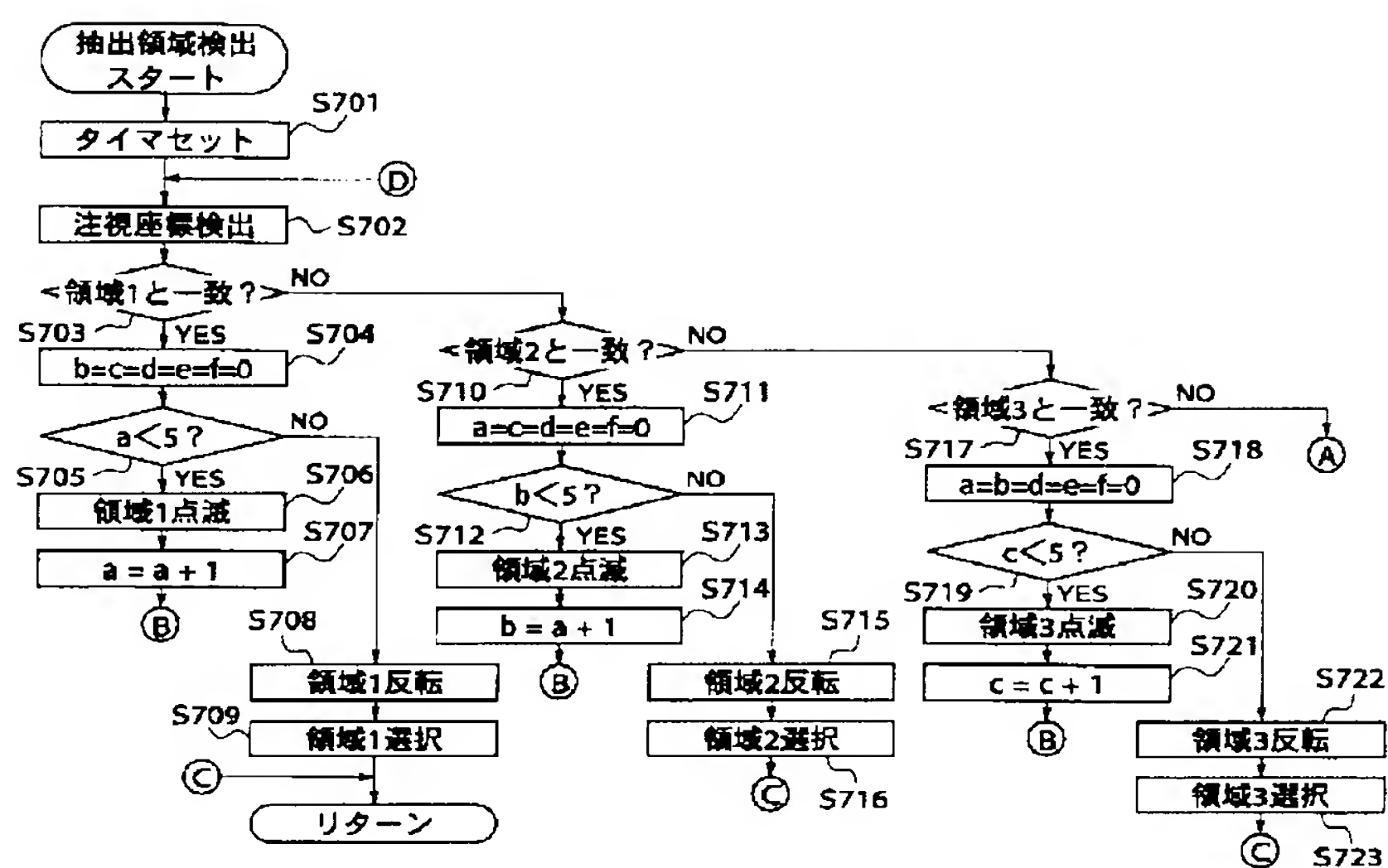
【図11】



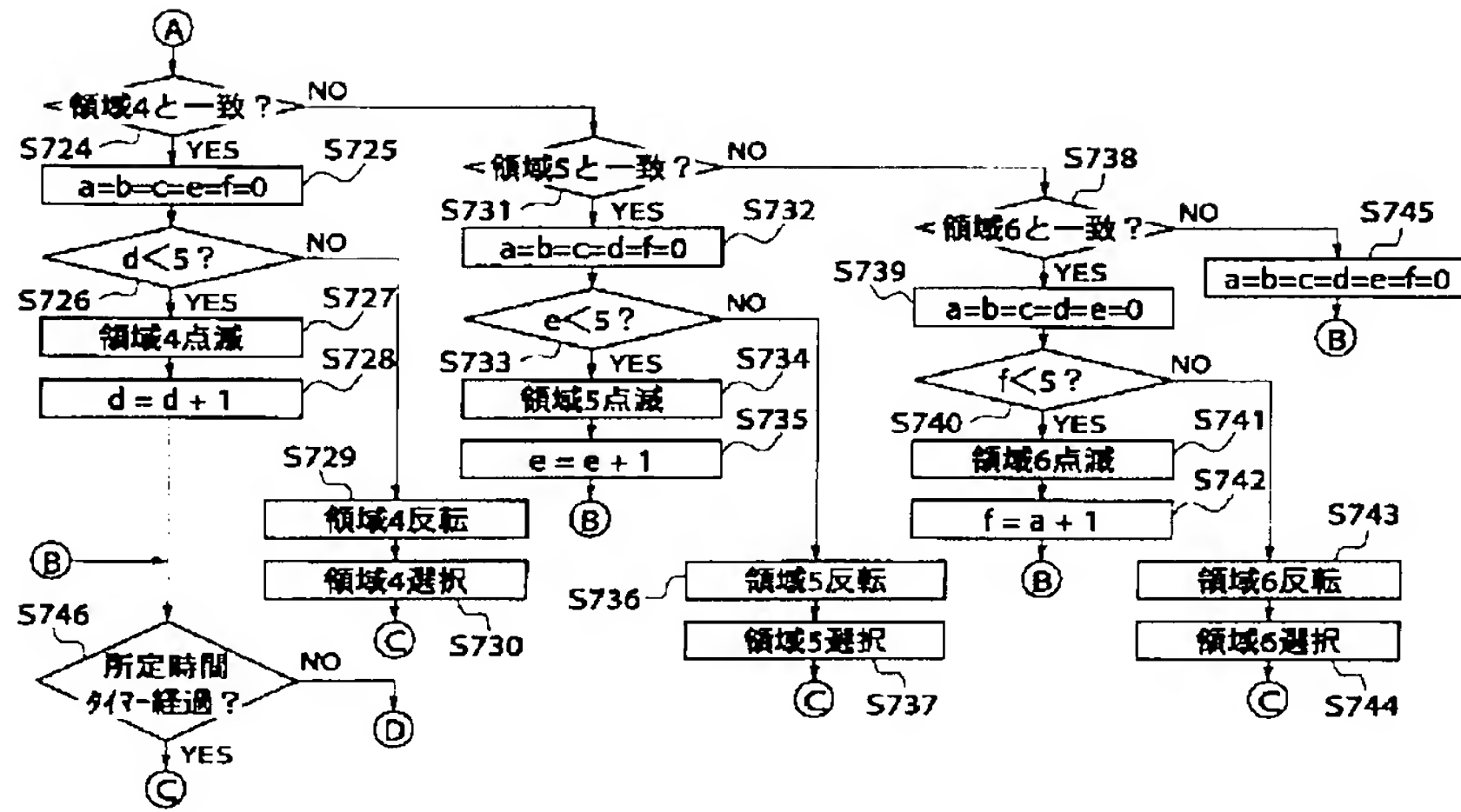
【図6】



【図7】



【図8】



【図12】

